CDF note 10593

A Search for the Standard Model Higgs Boson in the Process $ZH \to e^+e^-b\bar{b}$ Using 7.5 fb^{-1} of CDF II Data

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: July 18, 2011)

We present the results of a search for the standard model (SM) Higgs boson in the associated production process $p\bar{p} \to ZH \to e^+e^-b\bar{b}$. We analyze a sample of Tevatron $p\bar{p}$ collisions at $\sqrt{s}$=1.96 TeV, corresponding to an integrated luminosity of 7.5 fb^{-1}, and collected by the CDF II detector. In events with two electron (e) candidates and at least two energetic jets, one or more of which is identified as a bottom-quark (b) jet, we set 95% confidence level upper limits on the associated Higgs production cross section (σ_{ZH}) times the branching ratio (BR) of $H \to b\bar{b}$ for eleven Higgs boson mass hypotheses between 100 GeV/c^2 and 150 GeV/c^2. For a Higgs boson mass of 115 GeV/c^2, we observe (expect) an upper limit of 3.91 (5.79) times the SM value of $\sigma_{ZH} \times BR(H \to b\bar{b})$.

Preliminary Results for Summer 2011 Conferences

I. INTRODUCTION

The mechanism of spontaneous electroweak symmetry breaking (EWSB) has yet to be resolved in the standard model of particle physics (SM). In the most simple formulation, EWSB is achieved through the introduction of the theoretical Higgs field [1–3], and consequently predicts the existence of the unobserved Higgs boson. Precision electroweak measurements constrain the mass of the Higgs boson (M_H) to be less than 169 GeV/c^2 at the 95% confidence level [4]. Previous searches at LEP-2 and the Tevatron have excluded masses below 114 GeV/c^2 and between 158 GeV/c^2 [5] and 173 GeV/c^2 [6] respectively. A Higgs boson with a mass of 135 GeV/c^2 or less will decay to a bottom quark pair ($b\bar{b}$) approximately 80% of the time [7]. While the rate of associated ZH production is small ($\sim$ 0.1 pb), the decay products of the associated Z boson (here, e^+e^-) serve to distinguish $H \to b\bar{b}$ events from multi-jet production, making $ZH \to e^+e^-b\bar{b}$ a viable Tevatron search mode.

In this note, we present an updated search for $ZH \to e^+e^-b\bar{b}$, in which we expand upon the techniques of and improve the electron component ($Z \to ee$) of previous searches performed in 4.1 fb^{-1} [8] and 5.7 fb^{-1} [9]. This search introduces new multivariate electron selections, boosted decision tree (BDT) [10] event classification, and in combination with a search in $ZH \to \mu^+\mu^-b\bar{b}$ [11], tiered artificial neural network (NN) discrimination combining specialized NNs designed for ZH, top-quark pair ($t\bar{t}$), and b-quark isolation.

This note is structured as follows: Section II describes the initial (trigger) online event selection and the data sample considered. Section III details the selections used to identify a sample of candidate events consistent with the expected topology of the $ZH \to e^+e^-b\bar{b}$ process. The composition of our selected sample, simulation of ZH signal, and the formulation of our data model are discussed in Section IV. The training and construction of our NN and BDT event discriminants are detailed in Section V. The extraction of upper limits on the value of $\sigma_{ZH} \times BR(H \to b\bar{b})$ is presented in VII.

II. DATA SAMPLE & ONLINE EVENT SELECTION

The data used in this search were collected by the upgraded CDF II detector, between February 2002 and March of 2011, and correspond to 7.5 fb^{-1} of Tevatron $p\bar{p}$ collisions at $\sqrt{s}$=1.96 TeV. The CDF II detector is described in detail elsewhere [12]. CDF II records only those collision events that meet the criteria of a multi-level online event selection (trigger) system. In this search, we use data selected by any of three trigger algorithms:

- The Central High p_T [13] Electron Trigger Algorithm:
  - Events with at least one central ($|\eta| < 1.1$) track of $p_T \geq$9 Gev/c whose extrapolation coincides with a calorimeter deposit of $E_T \geq 18$ GeV are selected by this trigger algorithm.
- The Dual Calorimeter Deposit Trigger Algorithm:
  - Events with at least two calorimeter deposits of $E_T \geq 18$ GeV within $|\eta| < 3.6$ are selected by this trigger algorithm.
- The Z Candidate Trigger Algorithm:
  - Events with at least one calorimeter deposit of $E_T \geq 18$ GeV and a second calorimeter deposit of $E_T \geq 9$ GeV with a combined mass greater than 40 GeV/c^2 are selected by this trigger algorithm. This trigger algorithm has not been previously used in a ZH search.

The three triggers have (in combination) collected a total of $\sim$ 600 million events that are subjected to the selections described below.

III. EVENT SELECTION

Events passing initial trigger selection are subject to three additional layers of selections.

- A. Electron and Z boson Candidate Selections.
- B. Jet Selections.
- C. b-quark Jet Selections. (b tagging)

Events meeting the requirements (details below) of both selection layers A and B are assigned the classification "PreTag". The Pretag sample serves as a high-statistics model-validation sample. Events passing layers A, B, and C are classified as "Tag" level events. Tag level events form our analysis 'signal' sample and are used to set limits on the ZH production rate.

A. Electron and Z Boson Candidate Selections

To identify $Z \to e^+e^-$ candidates among the ~ 600 million events passing trigger selection, we begin by rejecting events that were recorded without operational calorimeter, tracker or silicon vertex detectors. We further remove events that are not well contained in the detector (*i.e.* those with a primary interaction vertex located more than 60 cm from the center of CDF II), and events that are likely to have originated from non-collison sources (primarily cosmic rays).

Next, we retain only those events with at least two electron candidates. The electron selections used in previous CDF ZH searches have been superseded by newly developed NN electron identification algorithms. We define three classes of electron candidates (based on reconstruction quality, and location within the detector) and have developed a NN specific to each. The three classes are:

- The 'Central' class consists of candidates with $|\eta| < 1.1$.
- The 'Tight Forward' class is made up of electron candidates with $|\eta| > 1.1$ matched to a PHOENIX[14] reconstructed track.
- The 'Loose Forward' class contains electron candidates with $2.8 > |\eta| > 1.1$ that are not associated with a PHOENIX track.

Using a combination of simulated electrons, simulated jets, and jets recorded in collision data, the three NNs were trained to discriminate between electrons, and jets that mimic the signature of an electron in the CDF calorimeter. Compared to the electron selections of previous searches, NN electron identification achieves a similar selection efficiency while reducing the fraction of Z boson candidates reconstructed from one or more jets misidentified as electrons (fakes) by a factor of five.

The class-appropriate NN algorithm is applied only to electron candidates meeting the preselection cuts listed in Table I. The NNs combine relevant calorimeter and tracker parameters to form a single numerical value for a given electron candidate. We reject all Central and Loose Forward candidates with a NN value lower than 0.3, and Tight Forward candidates with a NN value lower zero. Within a given class, we further designate electron candidates as 'High Score' or 'Low Score' as detailed in Table II.

Category	EmEt(GeV)	Had/Em
Central	>9	<0.125
Tight Forward	>9	<0.0625
Loose Forward	>9	< 0.125

TABLE I. The NN preselection cuts for the electron categories. EmEt is the transverse energy in the electromagnetic calorimeter. Had/Em is the ratio of the energy deposited in the hadronic calorimeter to that deposited in the electromagnetic calorimeter. A separate NN is used for each of the three categories.

Candidates Z bosons are reconstructed from electron-candidate pairs consisting of two High-Score candidates or one High-Score candidate and a second Low-Score candidate. We do not consider pairs consisting of two "Loose Forward" electrons. In addition, we form Z boson candidates using a combination of a High-Score Central electron and an electron candidate reconstructed primarily from tracking information. These "Crack Track" electrons were used in previous ZH searches, and are required to fall within uninstrumented regions of the calorimeter (*i.e.* the track is extrapolated to a region of poor calorimeter coverage). We impose an opposite charge requirement on all Z candidates when both constituent electrons have $|\eta| < 1.1$. Events without an identified Z candidate with a reconstructed mass between 76 and 106 GeV/c^2 are rejected; events meeting this criterion have passed selection layer 'A' and are subject to to further consideration.

B. Jet Selection

Candidate ZH events are required to have a topology consistent with that of the $ZH \to e^+e^-b\bar{b}$ process. Therefore, events meeting the requirements of selection layer 'A' are further required to contain two or more energetic jets. Jets

Clasification	High Score	Low Score
Central	NN value ≥ 0.75	$0.75 >$ NN value > 0.3
Tight Forward	NN value ≥ 0.5	$0.5 >$ NN value > 0.0
Loose Forward	NN value ≥ 0.6	$0.6 >$ NN value > 0.3

TABLE II. Designations (High/Low Score) within the three electron classifications.

are reconstructed using the JETCLU [15] clustering algorithm which merges calorimeter deposits to form jets with a radius of 0.4 in η-ϕ space. Jet energies are corrected for regional variations in detector response, calorimeter coverage, and the energy contribution from additional $p\overline{p}$ interactions. We retain events with at least two jets with $|\eta| \leq 2.0$. We require at least one jet with $E_T \geq 25$ GeV, and a second with $E_T \geq 15$ GeV. To improve the measurement of jet energies, we apply an additional correction to the jets (in a given event) with the highest and next-to-highest $E'_T s$. The correction factors are the output values of a [16] Neural Network derived function. The NN was used in previous CDF ZH searches, and modifies measured jet energies based on the separation between the jets and the transverse missing energy ($\not\!\!E_T$) [17] - the more aligned a jet is with the $E'_T s$, the more likely it is that the jet energy was mismeasured producing false $\not\!\!E_T$. The NN corrections improve the dijet mass (M_{jj}) resolution and improve the ability to isolate a potential ZH signal from background processes. Events with a Z candidate and two or more jets meeting the above criteria form the PreTag sample.

C. b-quark Jet Selection (b tagging)

To form our final analysis samples, we include only the subset of PreTag events that contain one or more jets likely to have originated from a b-quark. Within the CDF II detector, b-quarks hadronize before decaying producing a displaced vertex. Detection of a b-quark is referred to as b tagging. We use two b-tagging algorithms: SecVtx (secondary vertex) [18], and JetProb (the probability the jet originated from the primary vertex) [19]. We combine SecVtx and JetProb to form the following (exclusive) event classifications (listed in descending order of selection preference) :

- In the 'Double Tight Tag' category, we require events to contain at least two jets with a Tight SecVtx classification.
- In the 'Double Loose Tag' category, we require events to contain at least one jet with a Loose SecVtx classification and a second JetProb tagged jet.
- In the 'Single Tight Tag' category, we require events to contain exactly one jet with a Tight SecVtx classification [20].

The events forming the three b-tag categories make up the final analysis channels (*i.e.* they have passed the requirements of selection layers A, B, and C). They are subject to different systematic uncertainties, background compositions, and predicted ZH fractions, and are therefore maintained as separate analysis channels. However, they are simultaneously examined for ZH content and jointly used to set upper limits on $\sigma_{ZH} \times BR(H \rightarrow b\bar{b})$.

IV. DATA MODEL

Background processes possessing a detector signature similar to the signal are mainly those that contain two electrons and two (or more) jets in the final state. The dominant background is Z + jets, with Z + lights flavor jets (u, d, s) forming the major background component before b-tag requirements are imposed. Z + jets events are modeled using an ALPGEN v2.10 prime MC [21] with PYTHIA [22] 6.325 for showering. Signal, diboson (ZZ, WZ, WW), and $t\bar{t}$ processes are modeled with PYTHIA 6.216 MC. The $t\bar{t}$ simulation assumes a top mass of 172.5 GeV/c^2. Simulated samples enter the model with a normalization determined by the integrated luminosity (7.5 fb^{-1}) and by the process production rate. We normalize Z + jets samples using the ALPGEN leading order cross section with an additional K-factor of 1.4 to account for the difference between LO and NLO production rate. The assumed diboson, Z+jets, and $t\bar{t}$ production rates are summarized in Table III. The SM ZH production rates assumed in this search are listed in Table IV.

We apply several corrections to the normalization of simulated samples to improve the agreement between the observed data and the model. We correct the luminosity profile of the simulated samples to match that observed in data. We correct the energy ($\sim$1%) of observed leptons to ensure agreement with the energy distributions measured

Process	Generator	σ
Z+l.f.	ALPGEN+PYTHIA	4.66 fb to 2111 pb
$Z+c\bar{c}$	ALPGEN+PYTHIA	148.4 to 1512 fb
$Z+b\bar{b}$	ALPGEN+PYTHIA	53.9 to 715.4 fb
WW	PYTHIA	11.34 pb
WZ	PYTHIA	3.47 pb
ZZ	PYTHIA	3.62 pb
$t\bar{t}$	PYTHIA	7.04 pb

TABLE III. Assumed production rates used in normalization of model components. The Z+jets simulations includes a K-factor of 1.4 to account for the difference between NLO and LO calculations.

M_H (GeV/c^2)	σ(fb)	BR($H \to b\bar{b}$)
100	169.8	0.8033
105	145.9	0.7857
110	125.7	0.7590
115	103.9	0.7195
120	90.2	0.6649
125	78.5	0.5948
130	68.5	0.5118
135	60.0	0.4215
140	52.7	0.3304
145	46.3	0.2445
150	40.8	0.1671

TABLE IV. Signal production rates (σ) and branching ratios (BR). We include an additional factor of 0.10095 to account for BR(Z to charged leptons).

in data. In addition, we apply factors that correct for differences in lepton and b-jet reconstruction and selection efficiencies. To account for the selection efficiency of the CDF II trigger system, we employ multivariate trigger emulation. In an independent trigger sample, a NN is trained to reproduce the combined trigger decision of the three trigger algorithms detailed above. The output of this NN is applied to the model as an additional normalization factor on a per-event basis, and accurately reflects the probability of selection by the relevant triggers.

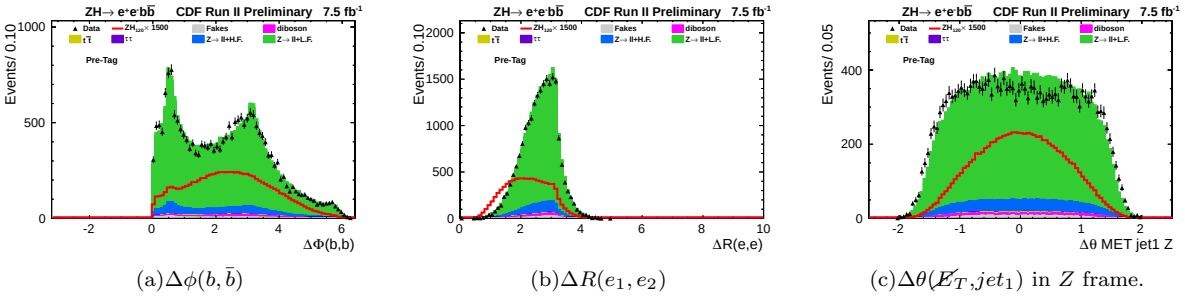
The probability of a jet to be misidentified as an electron is measured in independent jet-triggered data samples as a function of jet E_T and electron category. The rate is then applied to the jet plus electron pairs in data forming the "fake" $Z \to ee$ contribution. The fake events in the Crack Track plus central electrons category are modeled using same-sign events and added to the "fake" $Z \to ee$ sample.

The agreement between our model and observed data is checked across a variety of kinematic distributions, some of which are shown in Figures 1-7. Table V lists the predicted number of events from major backgrounds and the number of observed events at PreTag.

$ZH \to e^+e^-b\bar{b}$ Analysis CDF Run II Preliminary (7.5 fb^{-1})

Data	21122
ZH_{120}	6.2 ± 0.70
$t\bar{t}$	126 ± 17
Diboson	397 ± 34
$Z/\gamma^* \to ee$ + h.f.	1786 ± 561
$Z/\gamma^* \to ee$ + l.f.	18783 ± 4229
Fakes	354 ± 177
Model	21446 ± 4300

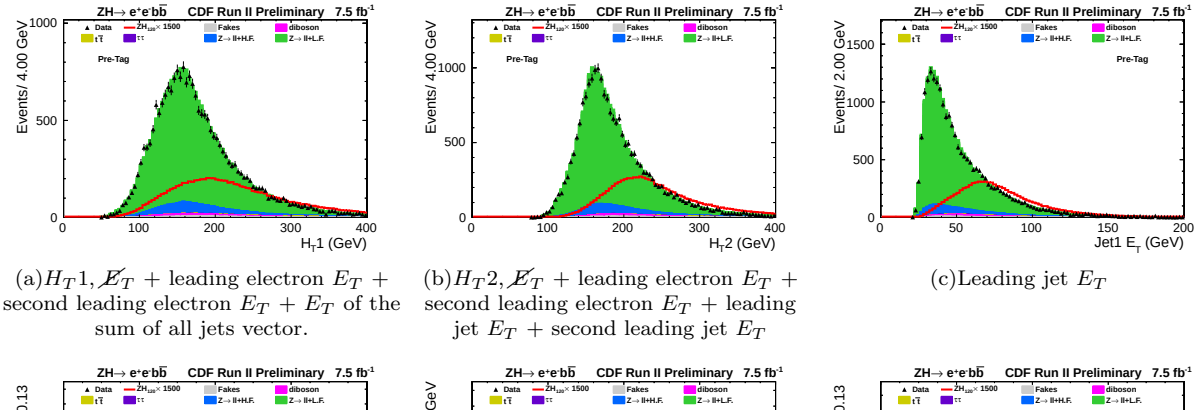
TABLE V. Event totals at the pretag level.



(a)$\Delta\phi(b,\bar{b})$ (b)$\Delta R(e_1,e_2)$ (c)$\Delta\theta(\not\!\!E_T,jet_1)$ in Z frame.

(d)$\Delta\theta$ $(\not\!\!E_T,jet_2)$ in Z frame. (e)$\Delta\theta$ $(\not\!\!E_T,e_1)$ in H frame. (f)$\Delta\theta(\not\!\!E_T,e_2)$ in H frame.

FIG. 1. Model plots at pretag.



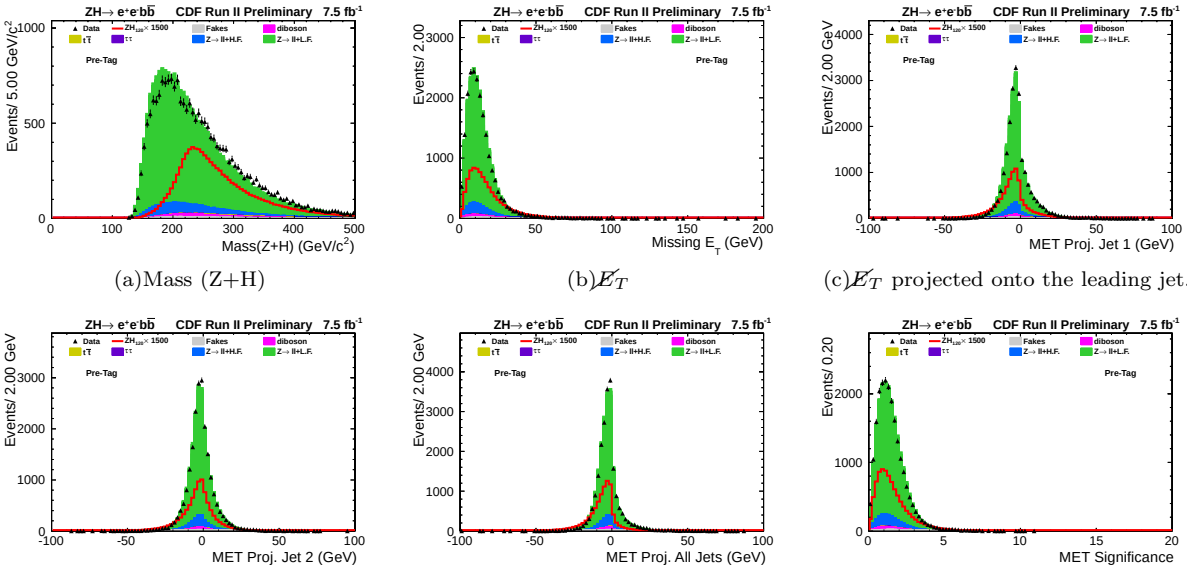
(a)$H_T1, \not\!\!E_T$ + leading electron E_T + second leading electron E_T + E_T of the sum of all jets vector.

(b)$H_T2, \not\!\!E_T$ + leading electron E_T + second leading electron E_T + leading jet E_T + second leading jet E_T

(c)Leading jet E_T

(d)Leading jet η (e)Second jet E_T (f)Second jet η

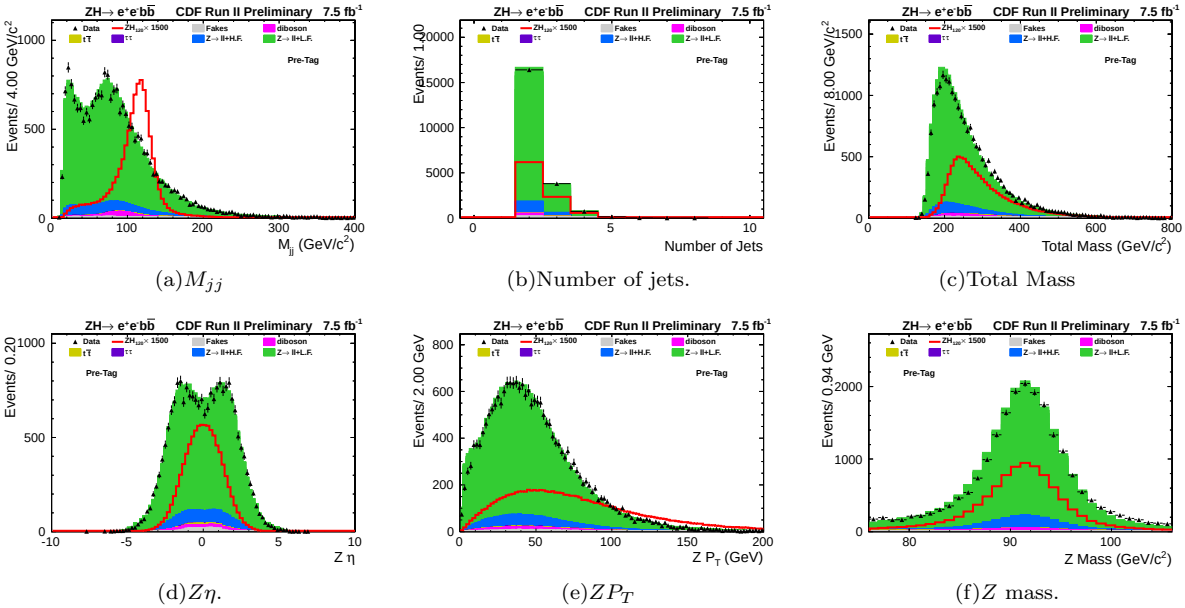
FIG. 2. Model plots at pretag.



(a)Mass (Z+H) (b)$\not\!\!E_T$ (c)$\not\!\!E_T$ projected onto the leading jet.

(d)$\not\!\!E_T$ projected onto the second leading jet. (e)$\not\!\!E_T$ projected onto the sum of all jets. (f)$\not\!\!E_T$ significance, $\not\!\!E_T/(j_1E_T + j_2E_T)$

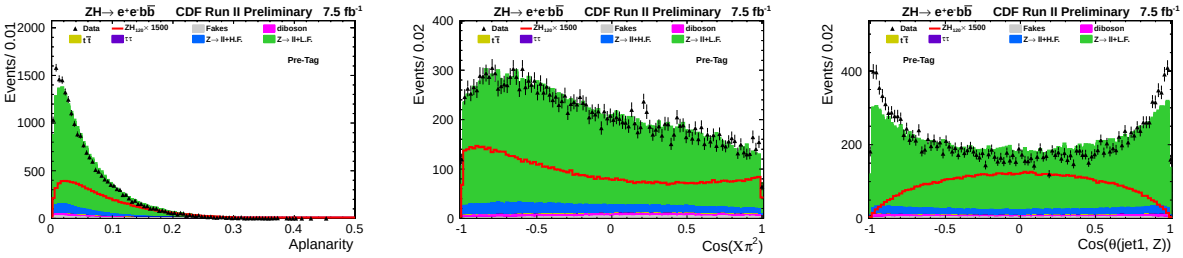
FIG. 3. Model plots at pretag.



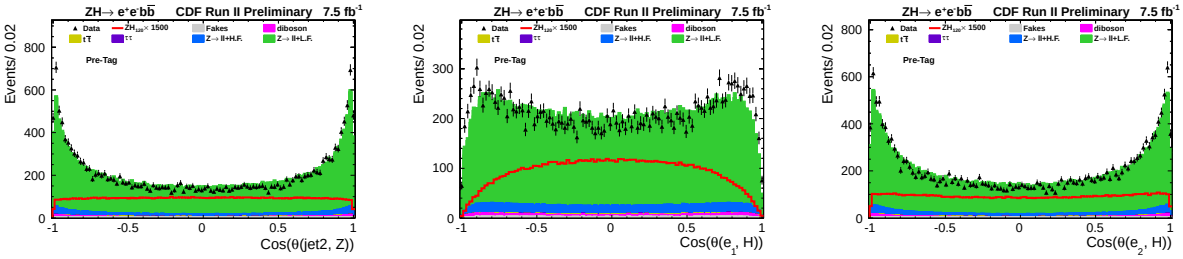
(a)M_{jj} (b)Number of jets. (c)Total Mass

(d)$Z\eta$. (e)ZP_T (f)Z mass.

FIG. 4. Model plots at pretag.

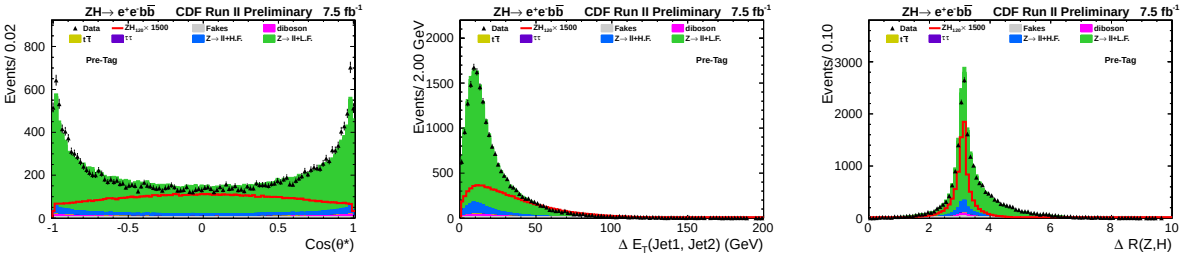


(a)Aplanarity. (b)Cos($\chi\pi^2$) (c)Cos($\theta(jet_1)$) in the Z frame.



(d)Cos($\theta(jet_2)$) in the Z frame. (e)Cos($\theta(e_1)$) in the H frame. (f)Cos($\theta(e_1)$) in the H frame.

FIG. 5. Model plots at pretag.

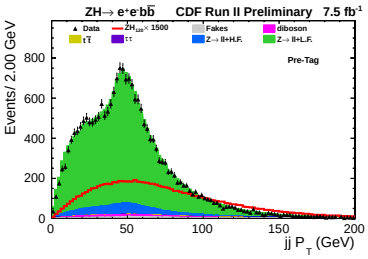


(a)Cos($\theta*$) (b)$\Delta E_T(j_1, j_2)$ (c)$\Delta R(Z, H)$

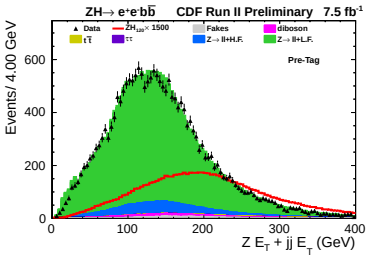
(d)$\Delta R(j_1, Z)$ (e)$\Delta R(j_1, j_2)$ (f)$\Delta R(j_2, Z)$

FIG. 6. Model plots at pretag.

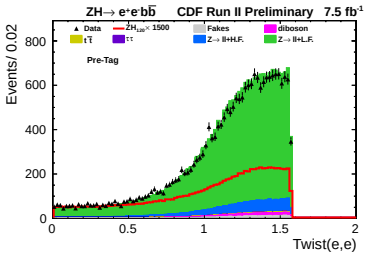
After b-tag selection, the contribution from Z + light flavor jets (mistagged jets) is modeled using re-weighted PreTag data events (*i.e.* we replace the ALPGEN Z+l.f. samples with a data-derived method), with the weights reflecting the probability for a light flavor jet to be tagged as a b-jet. These probabilities are measured for each b-tag algorithm in independent jet-triggered data samples.



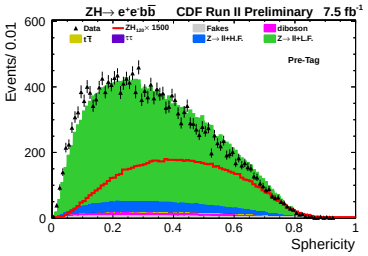
(a)HP_T



(b)ZE_T + Dijet E_T



(c)Twist(e, e)



(d)Sphericity

FIG. 7. Model plots at pretag.

V. EVENT DISCRIMINANTS

We use several multivariate algorithms to improve our ability to distinguish between ZH signal and background processes. To best use the information contained in a large number of distributions, we design two BDT's to isolate ZH events. One BDT (shape BDT) relies on event shape variables (*i.e.* the inputs summarize angular relationships, jet/lepton detector locations, *etc.*), while the other BDT (energy BDT) operates on distributions that quantify the energy content of the event (*i.e.* the inputs include jet energies, transverse momentums, *etc.*). The BDTs were trained to separate ZH events from the dominant tag-level backgrounds. The BDTs are used as input distributions to our final event discriminants and improve separation of ZH signal from background by $\sim 5\%$.

To form a final event discriminant, we use a series of artificial neural networks, and based on the individual NN scores, we separate the events into a binned final discriminant with three distinct regions (labeled I, II, and III) [11]. At Tag level the primary reducible backgrounds are Z+l.f./ jets and $t\bar{t}$. To isolate a potential ZH signal from these backgrounds we train a NN to distinguish between ZH signal and $t\bar{t}$ production, and apply a flavor separator based on the KIT [23] NN. KIT is designed to isolate incorrectly b-tagged jets from 'true' SecVtx-tagged b-quark jets. We form a variable based on KIT in combination with the JetProb b tagger, and define 'pseudo-KIT' (pKIT) as:

- the maximum of the KIT value of the jets in an event.
- one (1) if neither jet has a KIT score (KIT only returns real values for Tight SecVtx tagged jets), but one jet has a JetProb score <0.01.
- negative one (-1) if neither jet has a KIT score, and neither jet has a JetProb score <0.01.

A third NN (designated "Final NN") was developed to separate ZH from all backgrounds (Z+l.f. jets, Z+$b\bar{b}$, Z+$c\bar{c}$, diboson, and $t\bar{t}$) simultaneously. We employ eleven versions of this NN, each optimized for a different value of M_H. The inputs to these NNs are listed in Table VI. To assign events to one of the regions (I, II, III) we begin by assigning the event a score using the $t\bar{t}$ NN. If this score is greater than 0.5, the event is assigned to region I, otherwise the event is assigned an alternate score using the pKIT algorithm. If the pKIT score is less (greater) than zero, the events are assigned to region II (III). Once an event receives a region classification, it is evaluated by the Final NNs and assigned to a bin corresponding to this value. The logic of region and Final NN value assignment is outlined in Figure 8. The final discriminant outputs are shown at the PreTag level in Figure 9. The final Tag-level inputs are shown in Figures 10-25.

Final NN Discriminant Inputs
Cut Level
Energy BDT
Shape BDT
$\Delta R(e_1, e_2)$
Twist e_1e_2
Sphericity
$\Delta\phi(bb)$
Cos(θ^*)
$\Delta R(jet_2, Z)$
M_{jj}
$\not\!\!E_T$
Z.Et() + H.Et()
HP_T
ZP_T
MET projection onto the jets

TABLE VI. Distributions input to the final discriminants.
Twist(x_1, x_2)=$\tan^{-1}(\Delta\phi(x_1, x_2)/\Delta\eta(x_1, x_2))$ [24].
Cut Level is equal to -1 if PreTag or Single Tag, 0 if Loose Double Tag, and +1 if Tight Double Tag.
θ^* is the angle between the Z boson
candidate and the proton beam direction in the zero momentum frame.

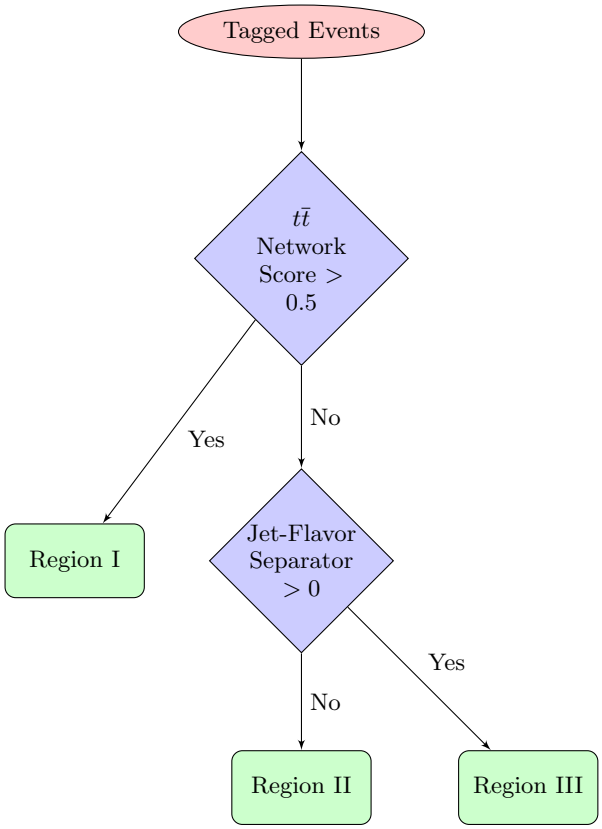
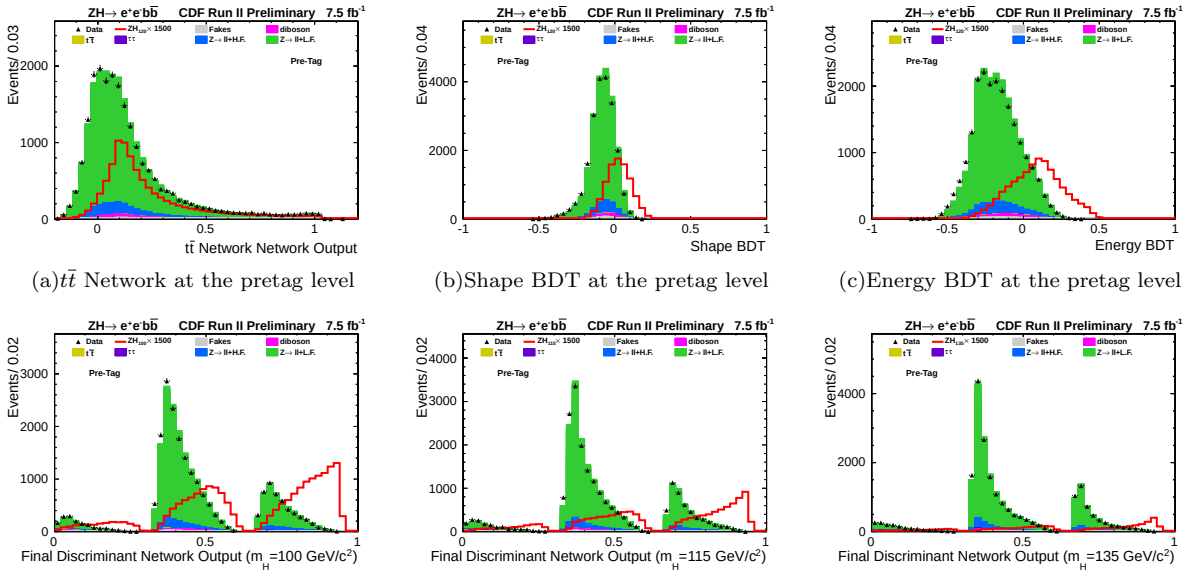


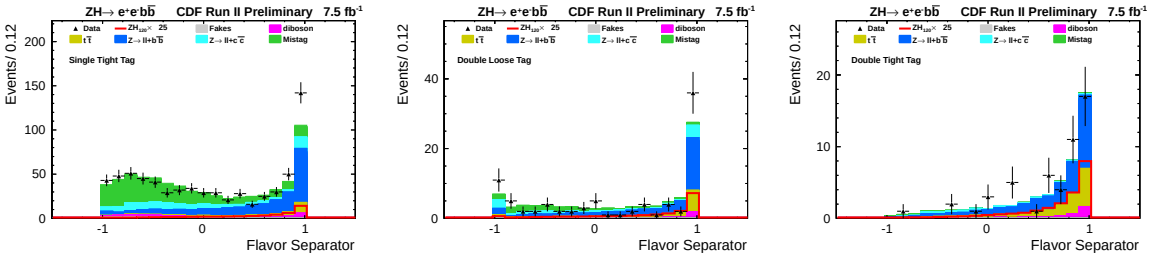
FIG. 8. Flow chart of how the final discriminant region bins are set. Within each of the three regions, a given event is assigned to the bin corresponding to the value output by the final neural network discriminant.



(a)$t\bar{t}$ Network at the pretag level (b)Shape BDT at the pretag level (c)Energy BDT at the pretag level

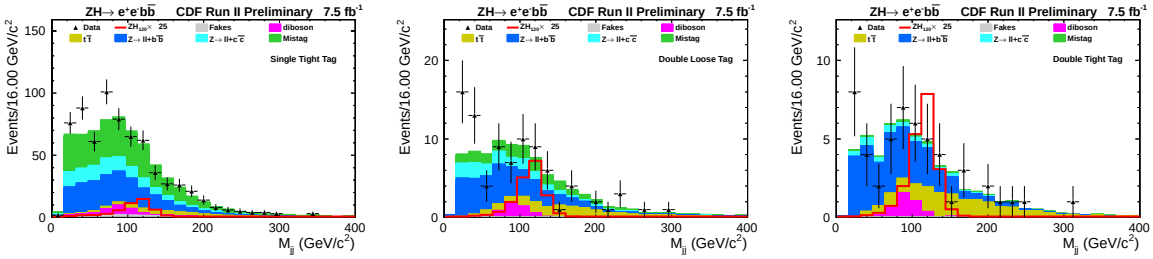
(d)ZH Network pretag level, m_H= 100 GeV/c^2 (e)ZH Network pretag level, m_H= 115 GeV/c^2 (f)ZH Network pretag level, m_H= 135 GeV/c^2

FIG. 9. Variables that determine the final discriminant region and final neural network outputs at the pretag level.



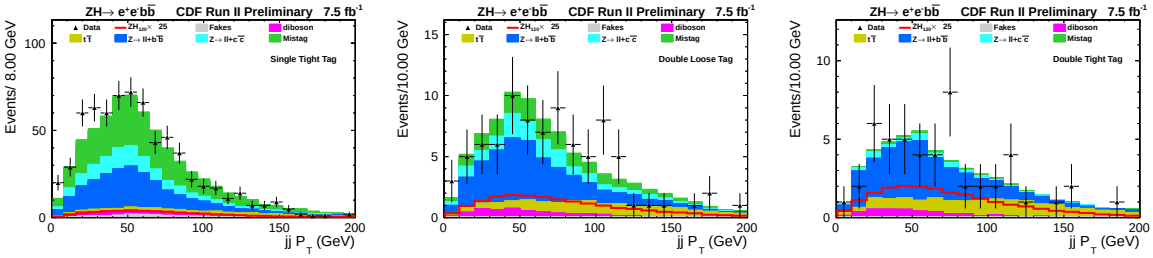
(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

FIG. 10. PKIT (flavor separator).



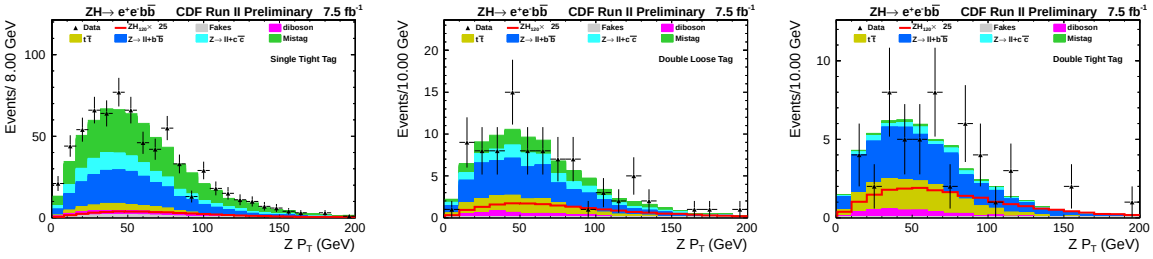
(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

FIG. 11. M_{jj}



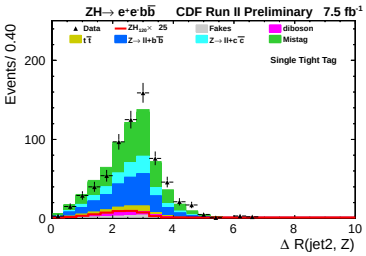
(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

FIG. 12. HP_T

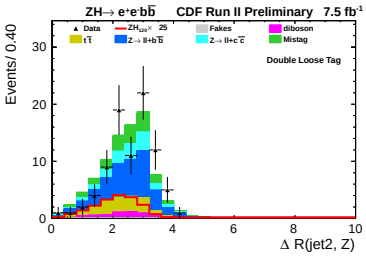


(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

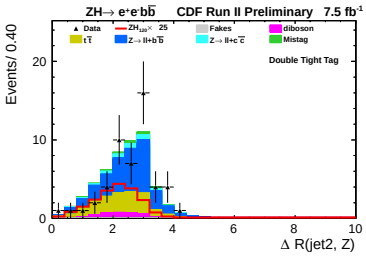
FIG. 13. ZP_T



(a)Single tight tag.

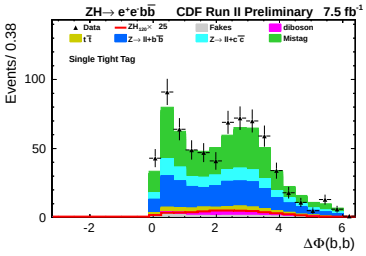


(b)Double loose tag.

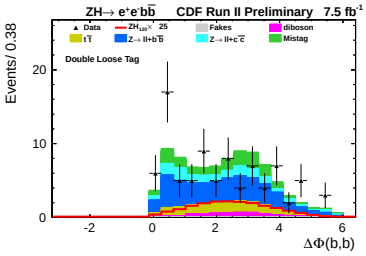


(c)Double tight tag.

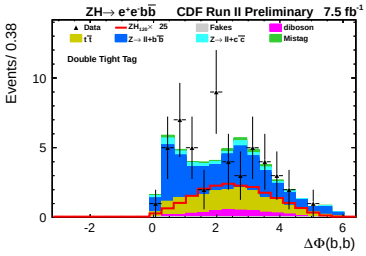
FIG. 14. $\Delta R(j_2, Z)$



(a)Single tight tag.

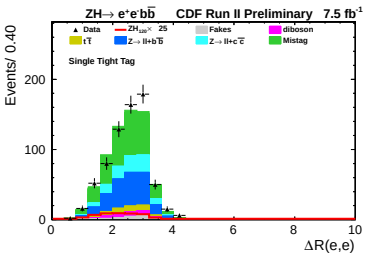


(b)Double loose tag.

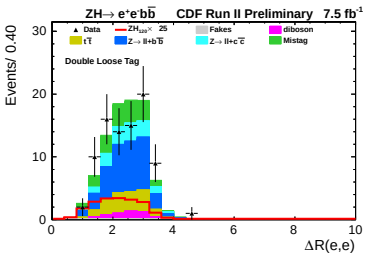


(c)Double tight tag.

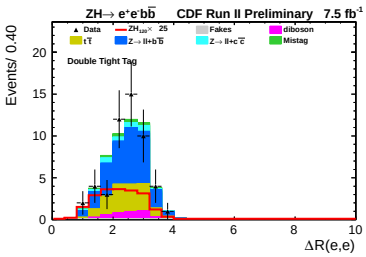
FIG. 15. $\Delta\phi(b, \bar{b})$



(a)Single tight tag.

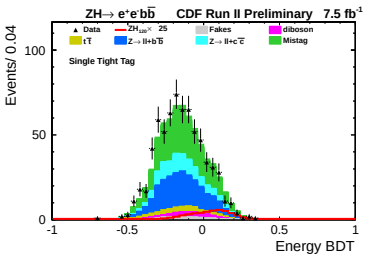


(b)Double loose tag.

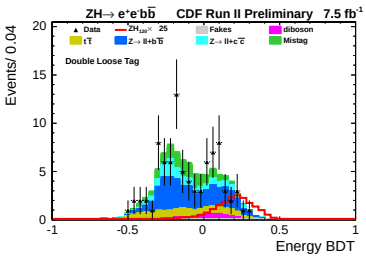


(c)Double tight tag.

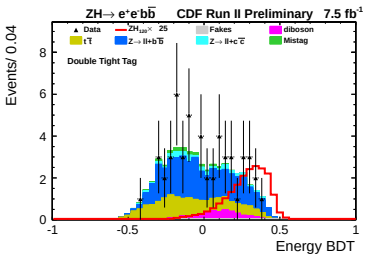
FIG. 16. $\Delta R(e_1, e_2)$



(a)Single tight tag.

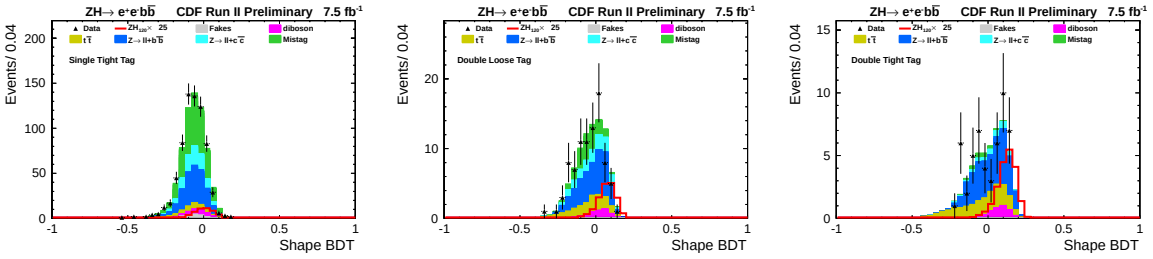


(b)Double loose tag.



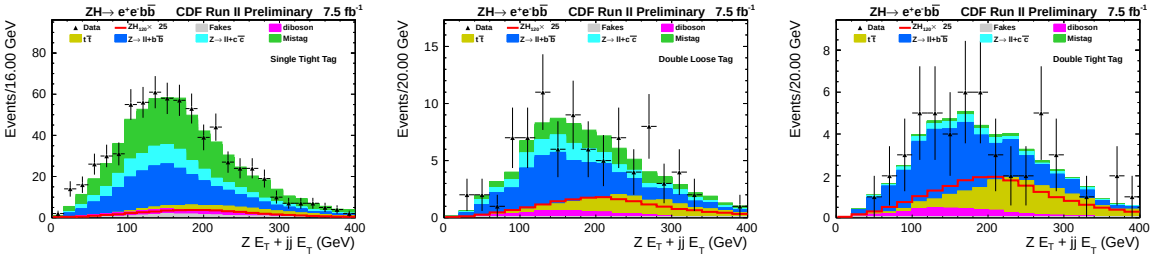
(c)Double tight tag.

FIG. 17. Energy boosted decision tree output (BDT).



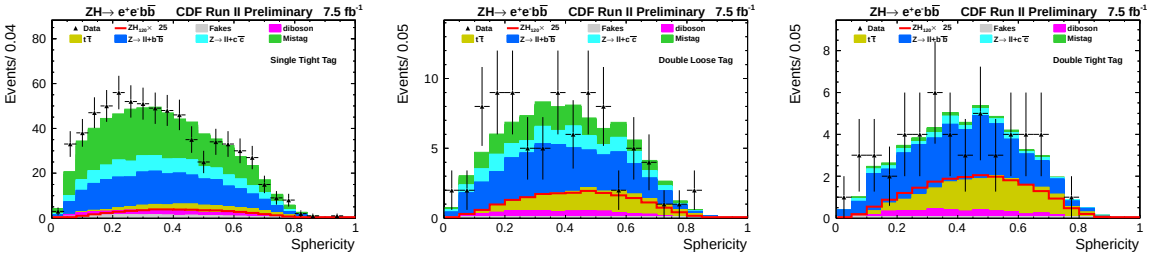
(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

FIG. 18. Shape boosted decision tree output (BDT).



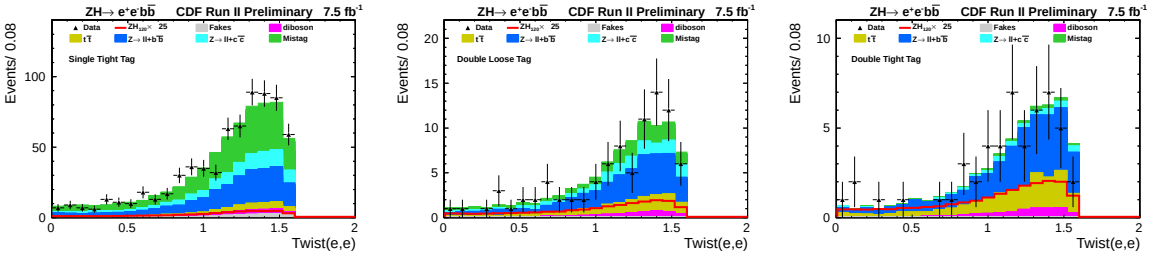
(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

FIG. 19. $ZE_T + jjE_T$



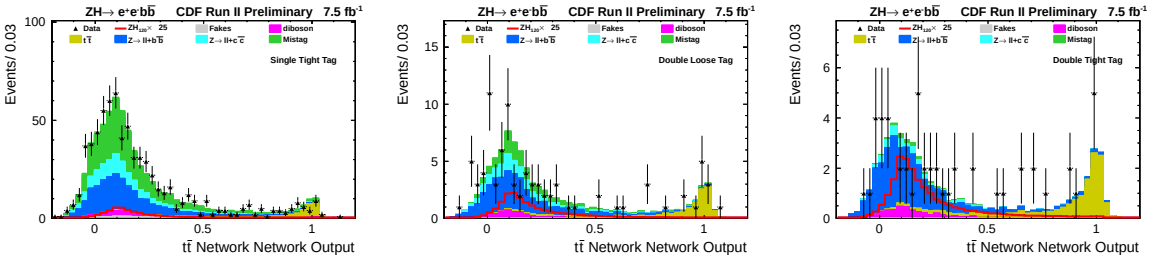
(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

FIG. 20. Sphericity.



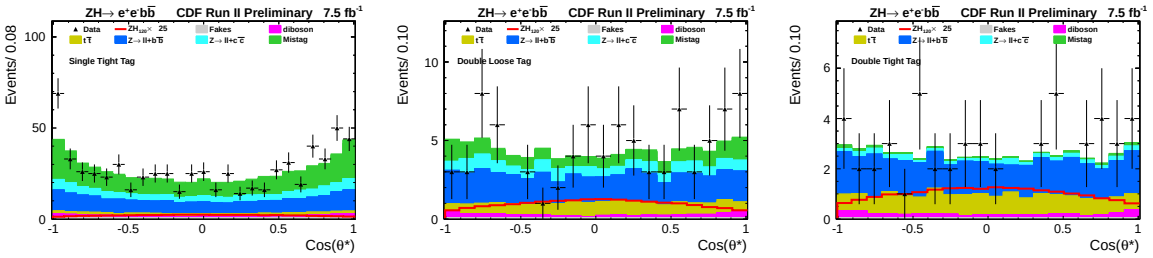
(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

FIG. 21. Twist(e_1, e_2)



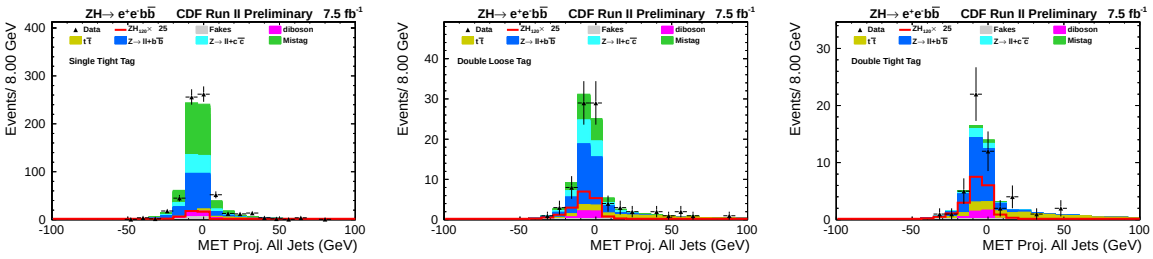
(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

FIG. 22. $t\bar{t}$ network output.



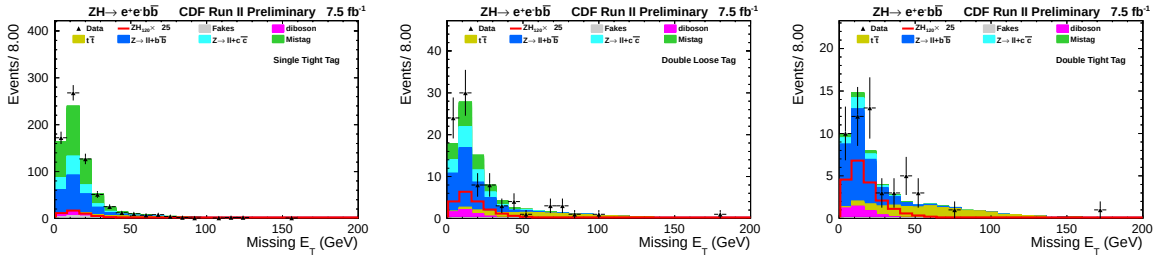
(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

FIG. 23. Cos($\theta*$).



(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

FIG. 24. $\not\!\!E_T$ projected on vector sum of all jets.



(a)Single tight tag. (b)Double loose tag. (c)Double tight tag.

FIG. 25. $\not\!\!E_T$

VI. SYSTEMATICS

We assume a ∼5% uncertainty on the integrated luminosity and we apply uncertainties of 5.2 (Single Tight Tag), 8.7 (Loose Double Tag), and 10.4 (Double Tight Tag)% to our b-tagged MC samples to account for the systematic errors associated with the calculation of b-tag efficiencies and scale factors. We apply 1% uncertainty on the trigger modeling. To account for systematics in our lepton ID scale factors and energy scales, we apply uncertainties of 2% and 3% respectively. In order to cover the spread of fake rates measured from different jet-triggered data samples, we assign a 50% uncertainty on our total fake estimate. We apply a 40% uncertainty to $Z + b\bar{b}$ and $Z + c\bar{c}$ samples to cover the theoretical uncertainty on the Z + heavy flavor jets cross-section. Similarly, we apply a 6% cross-section uncertainty to all diboson samples, and 5% to signal. For $t\bar{t}$, we place a 10% uncertainty to account for both the theoretical uncertainty on the process cross-section and the difference between our simulated top mass (172.5 GeV/c^2) and current experimental measurements. We assume an additional 4% rate uncertainty to account for the variation in ZH acceptance with different amounts of initial/final state radiation (ISR/FSR). We also consider uncertainties which are expected to affect both sample normalizations and the shape of the NN output distributions, such as the jet-energy scale (JES), and the mistag event weights. Tables VII through IX summarize the systematic uncertainties applied in our limit calculations.

Single Tight Tag $ZH \to e^+e^-b\bar{b}$ Analysis CDF Run II Preliminary (7.5 fb^{-1})

Contribution	Fakes	Top	WW	WZ	ZZ	$Z + b\bar{b}$	$Z + c\bar{c}$	Z+l.f.	ZH
Luminosity ($\sigma_{\text{inel}}(p\bar{p})$)	0	3.8	3.8	3.8	3.8	3.8	3.8	0	3.8
Luminosity Monitor	0	4.4	4.4	4.4	4.4	4.4	4.4	0	4.4
Trigger Emulation	0	1	1	1	1	1	1	0	1
Lepton ID	0	2	2	2	2	2	2	0	2
Lepton Energy Scale	0	3	3	3	3	3	3	0	3
ZH Cross Section	0	0	0	0	0	0	0	0	5
Fake Leptons	50	0	0	0	0	0	0	0	0
B-Tag Efficiency	0	5.2	5.2	5.2	5.2	5.2	5.2	0	5.2
$t\bar{t}$ Cross Section	0	10	0	0	0	0	0	0	0
Diboson Cross Section	0	0	6	6	6	0	0	0	0
$\sigma(p\bar{p} \to Z + HF)$	0	0	0	0	40	40	40	0	0
ISR/FSR	0	0	0	0	0	0	0	0	4.0
Mistag Rate (shape dep.)	0	0	0	0	0	0	0	$^{+13.9}_{-13.8}$	0
Jet Energy Scale (shape dep.)	0	$^{+1.9}_{-2.5}$	$^{+19.6}_{-4.0}$	$^{+5.2}_{-6.2}$	$^{+5.3}_{-7.1}$	$^{+12.1}_{-11.1}$	$^{+4.1}_{-9.9}$	0	$^{+3.0}_{-4.3}$

TABLE VII. Systematic uncertainties on the contributions for the Single Tight Tag channel. Systematic uncertainties are listed by name. Systematic uncertainties for ZH shown in this table are obtained for $m_H = 115$ GeV/c^2. Uncertainties are relative, in percent and are symmetric unless otherwise indicated.

Loose Double Tag $ZH \to e^+e^-b\bar{b}$ Analysis CDF Run II Preliminary (7.5 fb^{-1})

Contribution	Fakes	Top	WW	WZ	ZZ	$Z+b\bar{b}$	$Z+c\bar{c}$	Z+l.f.	ZH
Luminosity ($\sigma_{\rm inel}(p\bar{p})$)	0	3.8	3.8	3.8	3.8	3.8	3.8	0	3.8
Luminosity Monitor	0	4.4	4.4	4.4	4.4	4.4	4.4	0	4.4
Trigger Emulation	0	1	1	1	1	1	1	0	1
Lepton ID	0	2	2	2	2	2	2	0	2
Lepton Energy Scale	0	3	3	3	3	3	3	0	3
ZH Cross Section	0	0	0	0	0	0	0	0	5
Fake Leptons	50	0	0	0	0	0	0	0	0
B-Tag Efficiency	0	8.7	8.7	8.7	8.7	8.7	8.7	0	8.7
$t\bar{t}$ Cross Section	0	10	0	0	0	0	0	0	0
Diboson Cross Section	0	0	6	6	6	0	0	0	0
$\sigma(p\bar{p} \to Z+HF)$	0	0	0	0	40	40	40	0	0
ISR/FSR	0	0	0	0	0	0	0	0	4.0
Mistag Rate (shape dep.)	0	0	0	0	0	0	0	$^{+25.5}_{-21.4}$	0
Jet Energy Scale (shape dep.)	0	$^{+1.3}_{-2.3}$	0	$^{+7.5}_{-0.1}$	$^{+4.1}_{-4.4}$	$^{+8.2}_{-7.8}$	$^{+3.3}_{-5.5}$	0	$^{+2.1}_{-2.7}$

TABLE VIII. Systematic uncertainties on the contributions for the Loose Double Tag channel. Systematic uncertainties are listed by name. Systematic uncertainties for ZH shown in this table are obtained for $m_H = 115$ GeV/c^2. Uncertainties are relative, in percent and are symmetric unless otherwise indicated.

Tight Double Tag $ZH \to e^+e^-b\bar{b}$ Analysis CDF Run II Preliminary (7.5 fb^{-1})

Contribution	Fakes	Top	WZ	ZZ	$Z+b\bar{b}$	$Z+c\bar{c}$	Z+l.f.	ZH
Luminosity ($\sigma_{\rm inel}(p\bar{p})$)	0	3.8	3.8	3.8	3.8	3.8	0	3.8
Luminosity Monitor	0	4.4	4.4	4.4	4.4	4.4	0	4.4
Trigger Emulation	0	1	1	1	1	1	0	1
Lepton ID	0	2	2	2	2	2	0	2
Lepton Energy Scale	0	3	3	3	3	3	0	3
ZH Cross Section	0	0	0	0	0	0	0	5
Fake Leptons	50	0	0	0	0	0	0	0
B-Tag Efficiency	0	10.4	10.4	10.4	10.4	10.4	0	10.4
$t\bar{t}$ Cross Section	0	10	0	0	0	0	0	0
Diboson Cross Section	0	0	6	6	0	0	0	0
$\sigma(p\bar{p} \to Z+HF)$	0	0	0	40	40	40	0	0
ISR/FSR	0	0	0	0	0	0	0	4.0
Mistag Rate (shape dep.)	0	0	0	0	0	0	$^{+29.3}_{-25.4}$	0
Jet Energy Scale (shape dep.)	0	$^{+1.4}_{-2.6}$	$^{+7.8}_{-3.1}$	$^{+3.4}_{-5.9}$	$^{+6.8}_{-6.6}$	$^{+1.0}_{-3.7}$	0	$^{+1.6}_{-2.7}$

TABLE IX. Systematic uncertainties on the contributions for the Tight Double Tag channel. Systematic uncertainties are listed by name. Systematic uncertainties for ZH shown in this table are obtained for $m_H = 115$ GeV/c^2. Uncertainties are relative, in percent and are symmetric unless otherwise indicated.

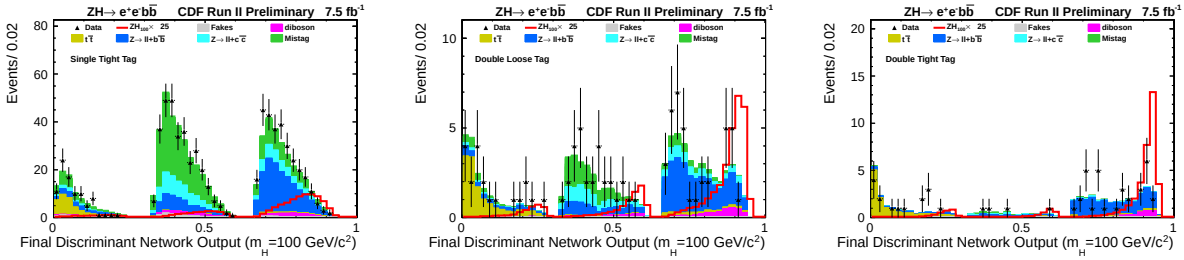
VII. RESULTS

After applying b tagging, our final event totals are shown in Table X. The neural net output distributions for the signal regions are shown in Figures 26-28. We do not observe a significant excess over the number of events predicted by our background model, and proceed to quantify the maximum allowed ZH contamination in the data. We use the MCLIMIT [25] machinery for this, and do a binned fit of the neural net distribution, including systematics. We set 95% confidence level upper limits on $\sigma_{ZH} \times BR(H \to b\bar{b})$ and compute observed limits for Higgs masses between 100 and 150 GeV/c^2 in 5 GeV intervals. The results are shown in Table XI, and Figure 30.

$ZH \to e^+e^-b\bar{b}$ Analysis CDF Run II Preliminary (7.5 fb^{-1})

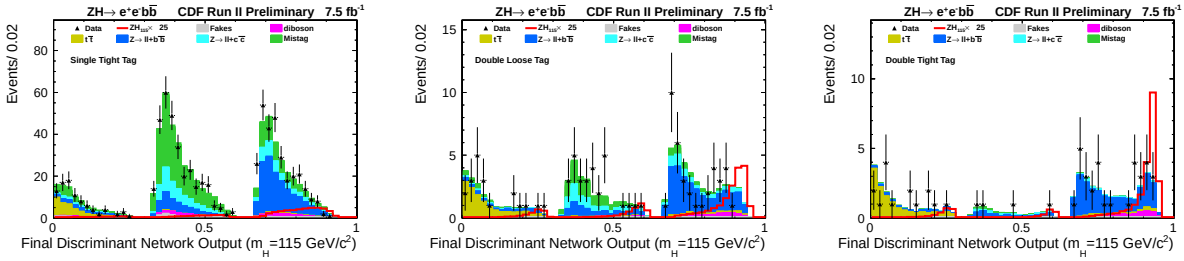
	Single Tight Tag	Loose Double Tag	Double Tight Tag
Data	693	87	51
ZH_{120}	2.0 ± 0.2	0.8 ± 0.1	0.9 ± 0.1
$t\bar{t}$	42 ± 6	17 ± 2	16 ± 3
Diboson	27 ± 3	5.7 ± 0.7	4.3 ± 0.6
$Z/\gamma^* \to ee$ + h.f.	254 ± 81	43 ± 14	27 ± 10
Mistags	333 ± 47	20 ± 5	2.2 ± 0.6
Fakes	25 ± 12	0.4 ± 0.2	0.2 ± 0.1
Model	681 ± 120	86 ± 20	50 ± 13

TABLE X. Event totals at the tag levels.



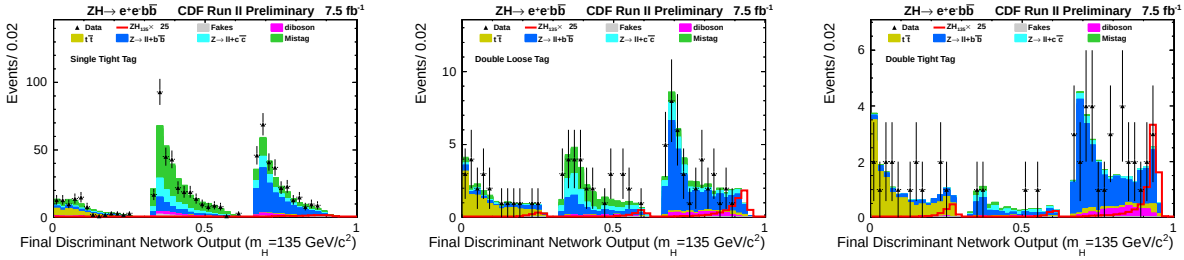
(a) (b) (c)

FIG. 26. Final discriminant for the m_H=100 GeV/c^2 mass point.



(a) (b) (c)

FIG. 27. Final discriminant for the m_H=115 GeV/c^2 mass point.



(a) (b) (c)

FIG. 28. Final discriminant for the m_H=135 GeV/c^2 mass point.

M_H (GeV/c^2)	Observed	-2σ	-1σ	Median	+1σ	+2σ
100	2.7	1.9	2.7	3.7	5.4	7.7
105	3.0	2.2	3.0	4.3	6.2	8.7
110	3.7	2.5	3.4	4.8	6.9	9.7
115	3.9	3.0	4.1	5.8	8.3	11.7
120	4.3	3.5	4.8	6.9	9.7	13.8
125	4.8	4.2	5.8	8.1	11.8	16.3
130	5.4	5.2	7.1	10.1	14.5	20.5
135	6.8	6.7	9.1	12.8	18.2	25.8
140	10.7	9.0	12.3	17.1	24.7	34.5
145	15.2	13.2	18.1	25.4	36.5	51.3
150	25.0	21.6	29.0	40.8	58.4	80.9

TABLE XI. The 95% CL upper limits on the ZH production rate expressed as a factor on $\sigma_{ZH} \times BR(H \to b\bar{b})$. The observed limits are obtained using CDF II data, while the median, -2,-1,+1, and +2 are obtained from the distribution of upper limits obtained in 5000 background-only pseudo-experiments.

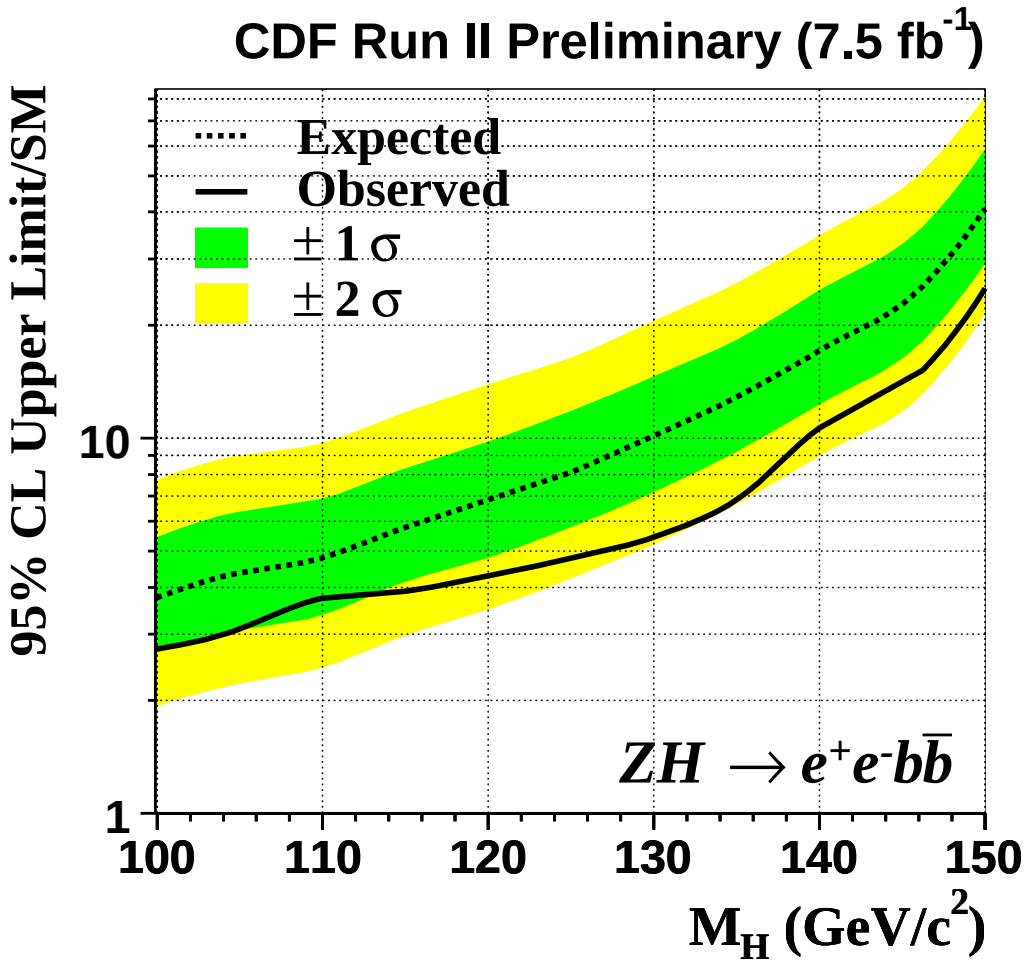
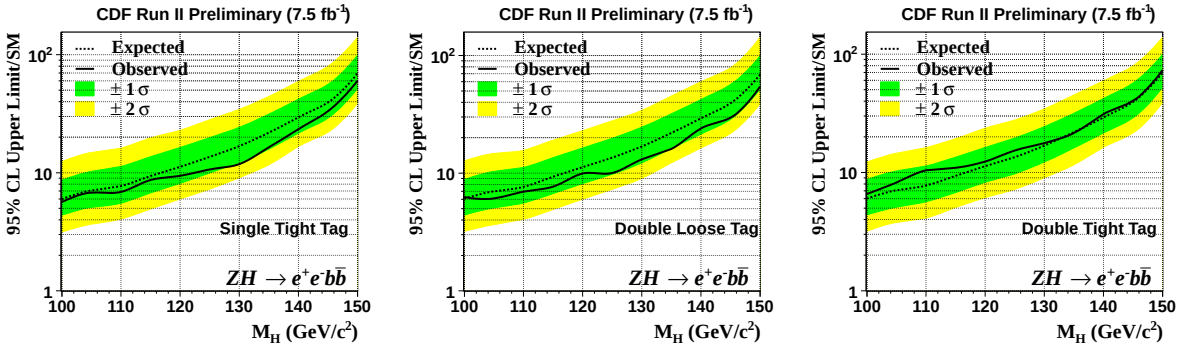


FIG. 29. The 95% CL upper limits on the ZH production rate expressed as a factor on$\sigma_{ZH} \times BR(H \to b\bar{b})$. The observed limits are obtained using CDF II data, while the median, -2,-1,+1, and +2 are obtained from the distribution of upper limits obtained in 5000 background-only pseudo-experiments.



(a)Single Tight Tag. (b)Double Loose Tag. (c)Double Tight Tag.

FIG. 30. The 95% CL upper limits on the ZH production rate expressed as a factor on$\sigma_{ZH} \times BR(H \to b\bar{b})$ in the individual tag channels. The observed limits are obtained using CDF II data, while the median, -2,-1,+1, and +2 are obtained from the distribution of upper limits obtained in 5000 background-only pseudo-experiments.

VIII. CONCLUSIONS

We have evaluated a new limit with an updated dataset 32% larger than in the previous analysis. We have calculated a 95% confidence level upper limits from 3.73 to 40.35 times the Standard Model prediction for Higgs boson masses between 100 GeV/c^2 to 150 GeV/c^2. For a Standard Model Higgs boson mass of 120 GeV, we find the expected 95% confidence level upper limit to be 6.92 times the Standard Model prediction with an observed limit of 4.21.

ACKNOWLEDGMENTS

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium fuer Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Comision Interministerial de Ciencia y Tecnologia, Spain; and in part by the European Community's Human Potential Programme under contract HPRN-CT-20002, Probe for New Physics.

[1] Higgs, P., Phys. Rev. Lett. **13** 508 (1964).
[2] Guralnik, G. S., *et. al.*, Phys. Rev. Lett. **13** 585 (1964).
[3] Englert, F. and R. Brout, Phys. Rev. Lett. **13** 321 (1964).
[4] Baak, M., *et. al.*, arXiv:1107.0975v1 [hep-ph] (2011)
[5] The LEP Electroweak Working Group. http://lepewwg.web.cern.ch/LEPEWWG/.
[6] CDF, D0, and the TEVNPHWG, arXiv:1103.3233v2 [hep-ex] (2011).
[7] Strange, A., *et. al.*, Phys. Rev. D **49** 1354 (1994).
[8] T. Aaltonen, (CDF Collaboration) Phys. Rev. Lett. **105** 251802 (2010).
[9] CDF Collaboration Note **10235**.
[10] http://tmva.sourceforge.net/#credits.
[11] CDF Collaboration Note **10572**.
[12] Acosta, D., (CDF Collaboration), Phys. Rev. D. **71**, 032001 (2005).
[13] We use a cylindrical coordinate system with z along the proton beam direction, r the perpendicular radius from the central axis of the detector, and ϕ the aximuthal angle. for θ, the polar angle from the proton beam, we define η = -ln tan (θ/2). P_T refers to transverse momentum, while E_T denotes transverse energy.
[14] CDF Collaboration Note **6278**. The PHOENIX algorithm is employed for forward $|\eta| > 1.1$ track reconstruction.

[15] Blazey, G.C. and Flaugher, B.L., Ann. Rev. Nucl. Part. Sci. **49** 633-685 (1999).
[16] http://schwind.home.cern.ch/schwind/MLPfit.html
[17] The missing E_T ($\vec{\not\!\!E}_T$) is defined by the sum over calorimeter towers: $\vec{\not\!\!E}_T = -\sum_i E_T^i \hat{n}_i$, where i = calorimeter tower number with $|\eta| < 3.6$, $\hat{n}_i$ is a unit vector perpendicular to the beam axis and pointing at the ith calorimeter tower. We also define $\not\!\!E_T = |\vec{\not\!\!E}_T|$.
[18] Harrington, R. D., FERMILAB-THESIS-2007-25.
[19] Abulencia, A., (CDF Collaboration), Phys. Rev. D **74** 072006 (2006).
[20] Here, 'Tight' and 'Loose' refer to operating points of the individual b-tag algorithms. The Tight (Loose) algorithm points have lower (higher) tag efficiencies and lower (higher) rates of jet misclassification.
[21] Mangano, M., *et al.*, J. High Energy Phys. **07** 001 (2003).
[22] Sjöstrand, T., *et al.*, Comput. Phys. Commun. **135** 238 2001.
[23] CDF Collaboration Note **7816**.
[24] Black, K., *et. al.*, J. High Energy Phys. **1104** 069 (2011).
[25] Junk, T., Nucl. Instrum. Meth. **A434** 435 (1999).